MODIFIED PLANT VIRUSES AS VECTORS OF HETEROLOGOUS PEPTIDES

This invention relates to the use of plant viruses as carriers (vectors) for the production or presentation of foreign peptides. More particularly, the invention relates to the genetic manipulation of viral nucleic acid by incorporation of foreign nucleic acid sequences which are expressed as peptides in the virus particle (virion).

Our patent application WO 92/18618 (which is incorporated herein by reference) describes the utilisation of plant viruses as vector systems for the expression of foreign nucleotide sequences, i.e. nucleotide sequences (RNA or DNA) which are not present in plant viruses, as found in Nature, and which in consequence code for peptides not normally found in any naturally occurring plant virus. The invention described therein comprises assembled particles of a plant virus containing a foreign peptide. The plant viruses described therein are therefore modified forms of the native viruses and for convenience will be referred to as modified viruses. The present invention concerns the same modified viruses.

The foreign peptide which may be incorporated into plant viruses according to our prior application WO 92/18618 may be of highly diverse types and are subject only to the limitation that the nature and size of the foreign peptide and the site at which it is placed in or on the virus particle do not interfere with the capacity of the modified virus to assemble when cultured *in vitro* or *in vivo*.). In this specification the term "foreign", as applied to a peptide or to the nucleic acid encoding it , signifies peptides or nucleic acid sequences which are not native to the plant virus used as a vector. Such sequences can be alternatively described as exogenous or heterologous sequences. The term "peptide" includes small peptides and polypeptides. The peptide preferably contains more than 5 amino acid residues. In broad concept, modified viruses may be formed from any biologically useful peptides (usually polypeptides) the function of which requires a particular conformation for its activity. This may be achieved by association of the peptide with a larger molecule, eg. to improve its stability or mode of presentation in a particular biological system. Examples of such peptides are peptide hormones; enzymes; growth factors; antigens of protozoal, viral, bacterial, fungal or animal origin; antibodies including anti-

idiotypic antibodies; immunoregulators and cytokines, eg interferons and interleukins; receptors; adhesins; and parts or precursors of any of the foregoing types of peptide.

Among the broad range of bioactive peptide sequences presented on plant virus vectors in accordance with WO 92/18618 special importance attaches to the antigenic peptides which are the basis of vaccines, particularly animal (including human) virus vaccines. It should be noted that in the context of WO 92/18618 vaccines may have prophylactic (i.e. disease prevention) or therapeutic (i.e. disease treatment) applications. For vaccine applications WO 92/18618 provides an especially attractive epitope presentation system. When used for such applications the antigenic peptide component will be sited appropriately on the virus particle so as to be easily recognised by the immune system, for example by location on an exposed part of the coat protein of the virus. As applied to the latter, therefore, WO 92/18618 preferably provides assembled particles of a modified plant virus containing an antigen derived from a pathogen, eg an animal virus, incorporated in an exposed position on the surface of the coat protein of the plant virus. The present invention also relates the use of such assembled modified plant virus particle as the immunogenic component of a vaccine. Such assembled modified plant virus particles presenting antigenic peptides also have applications as the antigen presentation component of an immunodiagnostic assay for detection of eg. animal (including human) pathogens and diseases.

The system described in WO 92/18618 is highly versatile in regard to the size of the foreign peptide which may be inserted into the viral coat protein. Thus peptides containing up to 38 or more amino acids have been successfully inserted in the course of our continuing research.

The first aspect of the present invention relates to assembled particles of a plant virus containing a predetermined foreign peptide as part of the coat protein of the virus.

However, the modified viruses so produced, being non-natural structures, are at a competitive disadvantage with the unmodified virus (wild type) when propagated in plants. As a result, we have observed a tendency in some modified viruses for the foreign peptide to be lost during

propagation with consequent reduction in yield of modified virus.

In accordance with our prior application WO 96/02649 (which is incorporated herein by reference), the causes of this problem have been identified and the steps necessary to avoid it have been determined.

Firstly, the process used for modifying the plant viral nucleic acid by introducing a nucleotide sequence coding for a foreign peptide should avoid the presence of direct sequence repeats flanking the insert. In the context of the present invention a construct containing a direct sequence repeat is one in which an identical oligonucleotide sequence is present on both sides of the inserted nucleotide. Such constructs can be genetically unstable because recombination can occur between the sequence repeats leading to loss of the foreign peptide coding sequence and reversion to the wild type sequence. Secondly, where the foreign oligonucleotide sequence is inserted into the plant virus genome as a substitution for part of the existing sequence, the resultant modified viral coat protein may be missing in an amino acid sequence which is important for virus replication, encapsidation and spread in the plant. This defect may be readily determined and avoided. Thirdly, the heterologous sequence should not be inserted at a sub-optimal site.

With reference to the above modification, a second aspect of the present invention relates to assembled particles of a plant virus containing a foreign peptide insert in the coat protein of the virus, the site of the insert in the coat protein corresponding to a site in the plant virus genome which is free from direct sequence repeats flanking the insert. Preferably the insert is an addition at a non-terminal site in the coat protein.

The present invention can be applied to any plant virus, including both icosahedral and non-icosahedral plant viruses. A preferred group of plant viruses for use as vectors are those in which the nucleic acid coding for the capsid is a separate moiety from that which codes for other functional molecules and whose coat proteins have a ß-barrel structure.

An advantage of the use of viruses which have a ß-barrel structure is that the loops between the individual strands of ß-sheet provide convenient sites for the insertion of foreign peptides. Modification of one or more loops is a preferred strategy for the expression of foreign peptides in accordance with the present invention.

To date, viruses from at least nine plant virus genera and three subgroup 2 ssRNA satellite viruses have had their tertiary and quaternary structures solved at high resolution. These are:

Table 1

Name	Acronym	Genus	Family
Southern bean mosaic virus	SBMV	*Sobemovirus*	not assigned
Sesbania mosaic virus	SMV	*Sobemovirus*	not assigned
Tomato bushy stunt virus	TBSV	*Tombusvirus*	*Tombusviridae*
Turnip crinkle virus	TCV	*Carmovirus*	*Tombusviridae*
Cowpea chlorotic mottle virus	CCMV	*Bromovirus*	*Bromoviridae*
Alfalfa mosaic virus	AMV	*Alfamovirus*	*Bromoviridae*
Bean pod mottle virus	BPMV	*Comovirus*	*Comoviridae*
Cowpea mosaic virus	CPMV	*Comovirus*	*Comoviridae*
red clover mottle virus	RCMV	*Comovirus*	*Comoviridae*
Tobacco ringspot virus	TRSV	*Nepovirus*	*Comoviridae*
Turnip yellow mosaic	TYMV	*Tymovirus*	not assigned
Tobacco necrosis virus	TNV	*Necrovirus*	*Tombusviridae*
satellite tobacco necrosis virus		*Subgroup 2*	
satellite panicum mosaic virus		*Subgroup 2*	
satellite tobacco mosaic virus		*Subgroup 2*	

All plant viruses possessing icosahedral symmetry whose structures have been solved conform to the eight stranded β-barrel fold as exemplified by cowpea mosaic virus, and it is likely that this represents a common structure in all icosahedral viruses. All such viruses are suitable for

use in this invention for the presentation of foreign peptide sequences, preferably in the loops between the ß-strands.

Preferred icosahedral plant viruses include all members of the following virus families: *Caulimoviridae, Bromoviridae, Comoviridae, Geminiviridae, Reoviridae, Partitiviridae, Sequiviridae, Tombusviridae*, and the following virus genera: Luteovirus, Marafivirus, Sobemovirus, Tymovirus, Enamovirus and Idaeovirus. Of the *Tombusviridae* family, the following genera are mentioned in particular: Dianthovirus, Machlomovirus and Necrovirus. An advantage of the *Comoviridae* and *Sequiviridae* is that their capsid contains sixty copies each of 3 different ß-barrels which can be individually manipulated. All other virus families and genera listed above have similar 3-dimensional structures but with a single type of ß-barrel. Viruses selected from the family *Comoviridae* (e.g. cowpea mosaic virus (CPMV), and bean pod mottle virus) are particularly preferred. CPMV is the most preferred virus.

In a particularly preferred embodiment the plant virus is cowpea mosaic virus (CPMV) and the foreign insert is made immediately preceding the proline 23 (Pro^{23}) residue in the ßB-ßC loop of the small capsid protein (VP23).

The present invention can also be applied to those β-barrel containing icosahedral plant viruses whose crystal structures have not yet been determined. Where significant sequence homology within the coat protein genes exists between one virus whose crystal structure is unknown and a second virus whose crystal structure has been determined, alignment of the primary structures will allow the locations of the loops between the β-strands to be inferred [see Dolja, V.V. and Koonin, E.V. (1991) J. Gen. Virol., 72, pp 1481-1486]. In addition, where a virus has only minimal coat protein sequence homology to those viruses whose crystal structure has been determined, primary structural alignments may be used in conjunction with appropriate secondary and tertiary structural prediction algorithms to allow determination of the location of potential insertion sites. The above application of the present invention is demonstrated in Examples 8 and 9.

A 3.5Å electron density map of CPMV (see Figure 1 in WO 92/18618) shows the clear

structural relationship between the capsids of CPMV and the T=3 plant viruses, for example the bromoviruses, in particular cowpea chlorotic mottle virus (CCMV) and the sobemoviruses, in particular southern bean mosaic virus (SBMV). The capsids of these latter viruses are composed of 180 identical coat protein subunits, each consisting of a single ß-barrel domain. These domains can occupy three different positions, namely A, B and C, within the virions (see Figure 1 in WO 92/18618). The two coat proteins of CPMV were shown to consist of three distinct ß-barrel domains, two being derived from the large capsid protein and one from the small capsid protein. Thus, in common with the T=3 viruses, each CPMV particle is made up of 180 ß-barrel structures. The single domain from the small subunit occupies a position analogous to that of the A type subunits of CCMV and SBMV and other viruses, whereas the N- and C-terminal domains of the large capsid protein occupy the positions of the C and B type subunits respectively (see Figure 1 in WO 92/18618).

X-ray diffraction analysis of crystals of CPMV and bean pod mottle virus (BPMV) shows that the 3-D structures of BPMV and CPMV are very similar and are typical of the *Comoviridae* in general.

In the structures of CPMV and BPMV, each ß-barrel consists principally of 8 strands of antiparallel ß-sheet connected by loops of varying length. The connectivity and nomenclature of the strands is given in Figure 2 of WO 92/18618. The flat ß-sheets are named the B,C,D,E,F,G,H and I sheets, and the connecting loops are referred to as the ßB-ßC, ßD-ßE, ßF-ßG and ßH-ßI loops.

One difference between the *Comoviridae* and the animal *Picornaviridae* is that the protein subunits of *Comoviridae* lack the large insertions between the strands of the ß-barrels found in *Picornaviridae*. The four loops (ßB-ßC, ßD-ßE, ßF-ßG and ßH-ßI - see Figure 3 in WO 92/18618) between the ß-sheets are suitable for expression of foreign peptides such as tumour-associated mucin peptide sequences.

The ßB-ßC loop in the small capsid protein is particularly preferred as the insertion site. This loop has an engineered *Aat*II site and a unique *Nhe*1 site at position 2708 of the M RNA-specific

sequence where foreign peptide sequences may be inserted (see Figure 4 of WO 92/18618). The insertion site immediately preceding Pro^{23} in the ßB-ßC loop of the small capsid protein is most preferred.

The present invention is equally applicable to non-icosahedral plant viruses and any of these viruses may be manipulated for the expression of peptides or polypeptides in accordance with the present invention. There are nine genera and one family of positive sense RNA rod-shaped plant viruses which do not possess icosahedral particle morphology. There are two types of rod shaped virus, rigid rods and flexuous rods. The genera containing rigid rods are: Tobamovirus, Tobravirus, Hordeivirus and Furovirus. The genera and family containing flexuous rods are: Potexvirus, Capillovirus, Trichovirus, Carlavirus, Closterovirus and Potyviridae. Preferred embodiments are: tobamoviruses, in particular tobacco mosaic virus (TMV) and sunn-hemp mosaic virus (SHMV); tobraviruses, in particular pea early browning virus (PEBV), pepper ringspot virus (PepRSV) and tobacco rattle virus (TRV); Potexviruses, in particular potato virus X (PVX), white clover mosaic virus (WCIMV) and clover yellow mosaic virus (CIYMV), Potyviruses, in particular potato virus Y (PVY), plum pox virus (PPV) and tobacco etch virus (TEV).

The structure of the tobamoviruses has been resolved to atomic resolution (by X-ray fibre diffraction) and it is to be assumed that the general architecture of the coat protein subunits of all rod shaped viruses are similar, differing only in the lengths and organisation of the N- and C-terminal extensions which protrude from the surface of the particle (Shukla *et al.* 1988). Application of the present invention to non-icosahedral plant viruses, the tertiary and quaternary structures of which are known to varying degrees, is demonstrated in Examples 10 to 13. As discussed above with reference to icosahedral plant viruses, the use of structural algorithms may be employed to identify potential insertion sites in non-icosahedral plant viruses whose crystal structures have not yet been determined.

To produce the modified plant virus particles in accordance with the first aspect of this invention the plant viral nucleic acid is modified by introducing a nucleotide sequence coding for the foreign peptide eg. an animal virus antigen at that part of the plant viral genome which codes for

the coat protein, infecting plants or plant cells with the modified viral nucleic acid, and harvesting assembled particles of the modified virus. Preferably, the nucleic acid sequence encoding the foreign peptide is introduced at the part of the plant virus genome which codes for an exposed portion of the coat protein. This procedure is best carried out by direct manipulation of the DNA of the virus in the case of DNA viruses or by manipulation of a cDNA corresponding to the RNA of an RNA virus. In the case of an RNA virus, an RNA transcript of the modified DNA is usually prepared for inoculation of plant cells or preferably whole plants so as to achieve a multiplication stage prior to the harvesting of assembled particles of the modified virus. Alternatively, cDNA clones of RNA viruses may be constructed in plasmids such that 5' ends of the viral coat protein encoding sequences are abutted directly to the transcriptional start site of a promotor active in the plant host (see Example 6). In the case of a DNA virus, the DNA itself is introduced into the plant. In this way, the foreign peptide is initially expressed as part of the capsid protein and is thereby produced as part of the whole virus particle. The peptide may thus be produced as a conjugate molecule intended for used as such. Alternately, the genetic modification of the virus may be designed in order to permit release of the desired peptide by the application of appropriate agents which will effect cleavage from the virus particle.

In order to produce modified virus on a commercial scale, it is not necessary to prepare ineffective inoculant (DNA or RNA transcript) for each batch of virus production. Instead, an initial inoculant may be used to infect plants and the resulting modified virus may be passaged in the plants to produce whole virus or viral RNA as inoculant for subsequent batches.

The foreign RNA or DNA may be inserted into the plant virus genome in a variety of configurations. For example, it may be inserted as an addition to the existing nucleic acid or as a substitution for part of the existing sequence, the choice being determined largely by the structure of the capsid portion and the ease with which additions or replacements can be made without interference with the capacity of the genetically-modified virus to assemble in plants. Determination of the permissible and most appropriate size of addition or deletion for the purposes of this invention may be achieved in each particular case by experiment in the light of the present disclosure. The use of additional inserts appears to offer more flexibility than replacement inserts in some instances.

In accordance with this invention, multiplication of modified virus and production of significant yields thereof in plant hosts is an important part of the novel strategy of the invention, in particular to produce antigens for vaccines and other types of peptide in an advantageous manner. As indicated above, the inserted heterologous nucleotide sequence may include those coding for amino acids which are readily cleaved so that, after a multiplication stage, the desired material may be separated from the virus particles. As an alternative to total cleavage of the peptide, it may be possible and desirable in some cases to release the peptide in a form in which it remains intact within a major part of the capsid but separated from the viral nucleic acid.

According to the second aspect of the present invention, two different restriction enzyme sites are chosen within the viral nucleic encoding the coat protein and the nucleic acid is cleaved using the appropriate restriction enzymes. Pairs of complementary oligonucleotides are synthesised encoding the foreign peptide which it is desired to insert into the virus coat protein. The oligonucleotides terminate in ends which are compatible with the restriction enzymes sites thus allowing insertion into the cleaved virus nucleic acid. This procedure results in the introduction of a nucleotide sequence coding for a foreign peptide whilst avoiding the presence of direct sequence repeats flanking the insert. Complementary oligonucleotides are synthesised in which the sequence encoding the heterologous amino acids are flanked by plant virus-specific sequences so that the foreign nucleic acid is inserted as an addition to the existing nucleic acid.

The coat proteins of a number of the viruses indicated in Table 1 has been compared. The similarity of the secondary structural elements and their spatial organisation is illustrated in Fig.10. Any of the loops which lie between the β-strands can be used for insertion of foreign epitopes, however the insertions are made such that the additions are exposed on either the internal or external surface of the virus and such that assembly of the coat protein subunits and the infectivity of the virus are not abolished. The choice of a particular loop can be made using knowledge of the structure of individual coat protein subunits and their interactions with each other, as indicated by the crystal structure, such that any insertions are unlikely to interfere with virus assembly. The choice of precise insertion site can be made, initially, by inspection of the crystal structure, followed by *in vivo* experimentation to identify the optimum site.

In a preferred embodiment the three dimensional structure of a plant virus is examined in order to identify portions of a coat protein which are particularly exposed on the virus surface and are therefore potentially optimum sites for insertion. In a further embodiment the amino acid sequence of the exposed portion of a coat protein is examined for amino acids which break α-helical structures because these are potentially optimum sites for insertion. Examples of suitable amino acids are proline and hydroxyproline, both of which whenever they occur in a polypeptide chain interrupt the α-helix and create a rigid kink or bend in the structure.

To demonstrate this system, the plant virus cowpea mosaic virus (CPMV) was chosen. The three-dimensional structure of CPMV has been solved at atomic resolution which has enabled identification of sites suitable for modification without disruption of the particle structure. CPMV comprises two subunits, the small (S) and the large (L) coat proteins, of which there are 60 copies of each per virus particle. Foreign peptide sequences may be expressed from either the L or S proteins or from both coat proteins on the same virion. Thus, up to 120 copies of the foreign peptide sequence may be expressed from a single virus particle.

Various sites in the CPMV coat protein have been identified as suitable for insertion of the foreign peptide. The co-ordinates given below refer to the linear amino acid sequence of the CPMV coat protein (S or L subunit).

Any insertion site which does not lie between the N-terminus of a subunit and a ß-strand, or between a ß-strand and the C-terminus, is considered to lie between two ß-strands. Such an insertion site may lie in a short loop at one of the axes of symmetry of the virus or in one of the much longer connecting strands which form the body of the protein subunits and which may contain additional secondary structure and form loops on the surface of the virus. In particular, there are α-helices present in some of the connecting strands which form the body of the protein subunits, and the co-ordinates given for some of the insertion sites may indicate that an α-helix is present between the insertion site and the preceding or succeeding ß-strand. For example, the S protein C' and C" ß-strands represent a secondary structure formed in the loop between the ßC and ßD strands.

(i) **External Surface Sites**

S Subunit (A Domain) Insertion Sites

ßB-ßC:

The residues between the ß-strands are Thr 19 to Val 22, and the preferred insertion site is between amino-acids 22 and 23. Insertion sites either side of this are also suitable, notably between residues 21 and 22.

ßC'-ßC":

The residues between the ß-strands are Val 42 to Asn 46.

ßH-ßI:

This site is at the tip of the five-fold axis and the residues between the ß-strands are Thr 152 to Gln 158.

ßD-ßE:

Again, this site is at the tip of the five-fold axis and the residues between the ß-strands are Ala 80 to Gln 90.

ßE-ßF:

This site is not at the tip of the five-fold axis, but lies 'behind' and to one side of the ß-strands. The residues between the ß-strands are Arg 96 to Ala 106. Residues 98 to 102 are preferred.

L Subunit, B Domain Insertion Sites

The B domain of the L subunit comprises amino acids 183-374 of the linear amino-acid sequence.

ßB-ßC:

This site is in the equivalent location on the subunit to the standard S protein insertion site and is at the three-fold axis of the virus. The residues between the ß-strands are Pro 201 to Glu 209.

ßH-ßI:

Again this site is at the three-fold axis of the virus and the residues between the ß-strands are His 331-Asp 341.

ßC-αA (ßC-ßD):

This site lies between the ßC and ßD strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αA) and the insertion site is a surface exposed portion which lies between the ßC strand and the αA helix. The surface exposed residues are Ala 223 to Ala 226.

ßG-αD (ßG-ßH):

This site lies between the ßG and ßH strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αD) and the insertion site is a surface portion which lies between the αD helix and ßH strands which are surface exposed. The surface exposed residues are Pro 314 to Thr 317.

ßE-αB (ßE-ßF):

This site lies between the ßE and ßF strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αB) and the insertion site is a surface portion which lies between the ßE strand and the αB helix. The surface exposed residues are Gly 269 to Phe 275.

L subunit, C Domain Insertion Sites

The C Domain of the L Subunit comprises amino-acids 1-182 of the linear amino-acid sequence.

ßE- αB (ßE-ßF):

This site lies between the ßE and ßF strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αB) and the insertion site is a surface exposed portion which lies between the ßE strand and the αB helix. The surface exposed residues are Gly 95 to Thr 102.

αD-ßH (ßG-ßH):

This site lies between the ßG and ßH strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αD) and the insertion site is a surface portion which lies between the αD helix and the ßH strands. The surface exposed residues are Ser 142 and Arg 145.

ßC-αA (ßC-ßD):

This site lies between the ßC and ßD strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αA) and the insertion site is a surface exposed portion which lies between the ßC strand and the αA helix. The surface exposed residues, not part of any secondary structural element, are Gly 53 to Phe 56.

ßB-ßC:

This site is an equivalent location on this domain to the S protein ßB-ßC (identified above) insertion site and is at the three-fold axis of the virus. The residues between the ß-strands are Ser 33 to Leu 42.

(ii) Internal Surface Sites

S Subunit (A Domain) Insertion Sites

ßG-ßH:

This protein chain between ß-strands points in towards the interior of the virus and forms a `double loop´. One insertion site comprises residues Pro 128 to Ser 130.

L Subunit B Domain Insertion Sites

ßF-ßG:

This loop is at the three-fold axis symmetry of the virus, and is the bottom loop of the four. The residues in the loop are Gln 287 to Glu 293.

C Domain ßl-B Domain ßB:

This is the junction between the B and C domains of the L subunit. This linker sequence comprises residues Asn 374 to Asp 186. The insertion site is around Ala 185, which is assigned to the B domain.

L Subunit, C Domain Insertion Sites

ßG-αD (ßG-ßH):

This site lies between the ßG and ßH strands. The protein chain loops out to form part of the body of the protein domain. Within this loop is an α-helix (termed αD) and the insertion site is an internal projecting loop which lies between the ßG strands and the αD helix. The residues in this loop are Asn 130 to Ser 135.

CPMV is biparite RNA virus and in order to manipulate the genome of any RNA virus to express foreign peptides it is necessary to use cDNA clones of the RNA. Full length cDNA clones of both CPMV RNA molecules are available which can be manipulated to insert oligonucleotide sequences encoding a foreign peptide. cDNA clones of the genome from plant RNA viruses can be used to generate *in vitro* transcripts that are infectious when inoculated onto plants. However, the infectivity of the transcripts is significantly lower than that of natural virion RNAs, probably as a result of the presence of non-viral residues at the termini of the transcripts. Difficulties may also be caused by exposure of the transcripts to degradative agents during inoculation. For this reason, the transcripts are usually stabilised by capping their 5' ends, but this is an inefficient, costly and time-consuming process.

In a further aspect of the present invention, cDNA clones of CPMV RNAs M and B have been constructed, in which the cDNA clone of the M RNA contains an inserted oligonucleotide sequence encoding a foreign peptide, which make use of the cauliflower mosaic virus (CaMV) 35S promoter sequence linked to the 5' ends of the viral cDNAs to generate infectious transcripts in the plant. This technique overcomes some of the problems encountered with the use of transcripts generated *in vitro* and is applicable to all plant RNA viruses.

To demonstrate the wide applicability of this invention, antigenic peptides from four different animal viruses, one bacterial pathogen of animals and a mammalian peptide hormone were used. Two of the viruses belong to the picornavirus group of animal viruses - foot and mouth disease virus (FMDV) and human rhinovirus (HRV). There are several important pathogens in this group, particularly, FMDV, poliomyelitis (polio) and hepatitis A. The third virus selected is human immune deficiency virus (HIV) which bears no similarity to any known plant virus, and for which no successful vaccines are currently available. The bacterial pathogen is *Staphylococcus aureus*, a causative agent of several animal diseases including mastitis in cows. The peptide hormone is porcine gonadotrophin releasing hormone.

The present invention will now be described with reference to the following Examples and accompanying drawings. No limitation thereto is intended:

Figure 1 depicts the region of CPMV M RNA which encodes the amino-terminal 40 amino acids of VP23. The numbers below the nucleotide refer to the M RNA sequence and the position of the unique *Nhe*1 site is indicated. The amino acids involved in forming the ßB and ßC strands of VP23 are indicated above the amino acid sequence of the protein which is shown using the standard one-letter code.

Figure 2 depicts (A) the sequence of the oligonucleotide used in the construction of pFMDV together with the amino acid sequence encoded by the top (positive) strand, which corresponds to amino acid residues 136-160 from VP1 of FMDV serotype O_1, and (B) the structure of VP23 after insertion of the FMDV-specific oligonucleotides. The arrowed region indicates the extent of the inserted FMDV epitope. The *Nhe*1 site not restored during the cloning is indicated by x*Nhe*1. The diagnostic *Bgl* II site present in the inserted sequence is also indicated.

Figure 3 depicts the effect of insertion of the FMDV-specific oligonucleotides, encoding amino acid residues 136-160 from VP1 of FMDV serotype O_1, on the structure of VP23 in pMT7-FMDV-I. The amino acids involved in forming the ßB and ßC strands of VP23 are indicated above the amino acid sequence of the protein which is shown using the standard one-letter code.

SEQ ID NOS:11 to 13

Figure 4 depicts the construction of a "substitution" vector by site-directed mutagenesis. The asterisk indicates the T residue that is changed to a C by site-directed mutagenesis, thereby creating a novel *Aat*II site.

SEQ ID NOS:14 to 16

Figure 5 depicts (A) the nucleotide sequence of the oligonucleotides used in the construction of pMT7-HIV together with the amino acid sequence encoded by the top (positive) strand, which corresponds to amino acid residues 735-752 from gp41 of HIV1, and (B) structure of VP23 after insertion of the HIV-specific oligonucleotides. The arrowed region indicates the extent of the inserted HIV epitope. The diagnostic *Pvu*1 site present in the inserted sequence is also indicated.

SEQ ID NOS:17 and 18

SEQ ID NOS:19 to 21

Figure 6 depicts (A) the nucleotide sequence of the oligonucleotides used in the construction of pMT7-HRV together with the amino acid sequence encoded by the top (positive) strand which corresponds to amino acid residues 85-99 from VP1 of HRV-14, and (B) the sequence of VP23 after insertion of the HRV-specific oligonucleotides. The arrowed region indicates the extent of the inserted HRV epitope. The diagnostic *Cla* 1 site present in the inserted sequence is also indicated.

SEQ ID NOS: 22 and 23

SEQ ID NO:24 to 26

Figure 7 depicts the effect of insertion of the FMDV-specific oligonucleotides, (depicted in bold type) encoding amino acid residues 141-160 from VP1 of FMDV serotype 0_1, on the sequence of VP23 in pMT7-FMDV-II.

SEQ ID NOS:27 to 41

Figure 8: Sequence of the oligonucleotides used to construct pMT7-FMDV, pMT7-HRV-II and pMT7-HIV-III. All oligonucleotides used terminated in the sequences shown in bold at the top of the diagram. The variable portions used for the construction of pMT7-FMDV-V (FMDV-V), pMT7-HRV-II (HRV-II) and pMT7-HIV-III (HIV-III) are shown below. The amino acid sequences encoded by the plus-sense oligonucleotides are indicated above the nucleotide sequence and correspond as follows: FMDV-V, amino acids 141-160 from VP1 of FMDV serotype 0_1; HRV-II amino acids 85-98 from VP1 of HRV-14; HIV-III, amino acids 731-752 from gp41 of HIV-1.

SEQ ID NO:33

SEQ ID NO:36

SEQ ID NO:39

Figure 9 Neutralization of HIV-1 IIIB by sera from individual C57/BL6 mice given two subcutaneous injections of the CPMV-HIV-I chimaera expressing amino acids 731-752 of gp41 on its surface (open bars). Mice were bled after 14 days. Also shown is the mean serum neutralization titre of a parallel group of mice inoculated with wild type CPMV (solid bars). All immunogens were formulated in alum adjuvant.

Fig. 10 is a simple line drawing of the solved β-barrel containing virus structures showing the secondary structural elements which make up the coat protein domains.

SEQ ID NOS: 42 to 58

Fig. 11 shows the nucleotide and protein sequences of SBMV surrounding a potential insertion site.

SEQ ID NOS: 59 to 61

FIG. 12 shows a comparison of ßH-ßI loop of three sobemoviruses. Conserved residues are highlighted in bold and the locations of the loops and ß-strands are indicated.

SEQ ID NOS: 62 to 80

Fig. 13 shows the nucleotide and protein sequences of LTSV surrounding a potential insertion site.

SEQ ID NOS: 81 and 82

Fig. 14 illustrates alignment of the coat protein sequences of RCNMV and TBSV using a Lipman-Pearson alignment algorithm.

Fig 15 illustrates a Chou-Fasman ß-region prediction plot of RCNMV residues 214-257 using an algorithm based upon the structures found in 64 proteins.

SEQ ID NO: 83

Fig. 16 illustrates application of the EMBL PHDsec algorithm program to the same RCNMV sequence as shown in Fig 15.

SEQ ID NOS: 85 TO 103

Fig. 17 shows the nucleotide and protein sequences of RCNMV surrounding a potential insertion site.

SEQ ID NOS: 104 to 120

Fig. 18 shows five deletion constructs and an unmodified clone of TRV as described in Example

13.

Modification of CPMV

Methods for manipulating the genome of the virus in order to make insertions into the coat of CPMV are described in WO 92/18618 and in WO 96/02649. A full length cDNA clone of CPMV M RNA in transcription vector pPMI is available (pPMM2902), as is a full length cDNA clone of CPMV (pBT7-123). A mixture of transcripts from pPMM2902 and pBT7-123 gives rise to a full virus infection when electroporated into cowpea protoplasts.

We have selected the ßB-ßC loop in VP23 for the insertion of foreign peptide. This loop is clearly exposed on the surface of the viral particle and computer modelling has shown that even large loops inserted at this site are unlikely to interfere with the interaction between adjacent subunits responsible for capsid structure and stability. This loop has a unique *Nhe*I site at position 2708 M RNA-specific sequence where foreign sequences may be inserted (see Figure 1).

The principal antigenic sites of the picornavirus foot and mouth disease (FMDV) and human rhinovirus (HRV), and the lentiretrovirus human immune deficiency virus (HIV) were used to illustrate the use of modified plant viruses in the production of vaccines to animal viruses.

The design and construction of pFMDV, a full length cDNA clone of CPMV M RNA containing an insert coding for a segment of FMDV loop protein, is described in WO 92/18618. An oligonucleotide sequence encoding amino acid residues 136-160 from VP1 of FMDV serotype O_1 strain BFS 1860 was inserted into the unique *Nhe*I site of pPMM2902 as an addition to the existing nucleic acid. The procedure used resulted in the creation of a direct repeat sequence flanking the insert (see Figure 2B). The properties of pFMDV transcripts are described in WO 92/18618. Infection of cowpea protoplasts with a mixture of pFMDV and pBT7-123 transcripts leads to multiplication and assembly of modified virus particles.

However, to produce modified plant viruses on a large scale it is necessary to prepare a construct

which can be inoculated directly onto whole plants, and which will replicate and assemble into virus particles as in the protoplast system. Therefore pPMM2902 was modified such that RNA synthesis is driven by a more efficient promoter and the modified plasmid transcribed under conditions that result in the transcripts having a "cap" structure at their 5' ends. The steps in the modification of pPMM2902 to produce pMT7-601 are described in detail in W0 92/18618. A mixture of capped pMT7-601 and pBT7-123 transcripts was found to be infectious to intact cowpea plants.

The design and construction of pMT7-FMDV-I, starting from pMT7-601 and pFMDV, are described in WO 92/18618. An oligonucleotide sequence encoding amino acid residues 136-160 VP1 of FMDV serotype 0_1, was inserted into the unique *Nhe*I site of pMT7-601 as an addition to the existing nucleic acid. The procedure used resulted in the creation of a direct repeat sequence flanking the insert (see Figure 3). The properties of pMT7-FMDV-I transcripts are described in detail in Usha, *et al.* [Virology (1993) 197, 366-3741]. Plants inoculated with a mixture of pMT7-FMDV-I and pBT7-123 transcripts developed lesions on their inoculated leaves which were smaller than those seen on the leaves of plants inoculated with wild type transcripts. Immunosorbent electron microscopy on leaf homogenates from inoculated leaves of pMT7-FMDV-I-infected plants confirmed the presence of CPMV-like virus particles. However, there was no evidence of systemic spread of the chimaeric virus particles to uninoculated leaves.

We have since characterised the progeny of a pMT7-FMDV-1 infection by reverse transcriptase-polymerase chain reaction (RT-PCR) analysis of RNA extracted from leaves inoculated with pMT7-FMDV-1. The analysis revealed the presence of two products, the major one corresponding to the expected product of approximately 580bp and a minor one of 500bp. The latter product comigrated with the product synthesised from RNA extracted from a plant infected wild-type CPMV. When the PCR products were cloned into bacteriophage M13 and the sequence around the site of insertion was determined, two classes of clones could be found: those which retained the entire FMDV-specific sequence (the majority) and those which contained a sequence corresponding exactly to wild-type CPMV (the minority). These results indicate that reversion to the wild-type sequence occurs in the transcript-inoculated leaves by an apparently single-step

process.

When RNA extracted from a pMT7-FMDV-1 transcript-inoculated leaf was passaged on to uninfected cowpeas, the plants developed symptoms on their inoculated leaves which consisted of a mixture of small lesions characteristic of a pMT7-FMDV-1 infection and larger wild-type CPMV lesions. In addition, the upper leaves developed mosaic symptoms characteristic of a wild-type CPMV infection. RT-PCR analysis of RNA extracted from the inoculated leaves of such again yielded two products but in this case the dominant one corresponded to that derived from wild-type M RNA. Analysis of clones derived from the dominant mixture of PCR products again revealed the same two classes of sequence found previously. However, in this case the majority of clones represented the wild-type sequence. These results indicate that not only does pMT7-FMDV-1 tend to lose the entire FMDV-specific sequence in a single step process, probably as a result of the presence of direct repeats flanking the insert but also that wild-type progeny have a significant advantage over the chimaeric virus.

In order to avoid the creation of a direct repeat sequence flanking the insert, a second restriction enzyme cutting site was created in the nucleotide sequence of the region of the CPMV genome encoding VP23. A single silent base change (U to C) at position 2740 of the M RNA creates a unique *Aat*II site at amino acid valine 27 (position 2735 of the nucleotide sequence). This was achieved by site-directed metagenesis of M13-JR-1 using methods described in WO 92/18618 (see Figure 4). The creation of the *Aat*II site enables the nucleotide sequence encoding the six amino acids from the native ßB-ßC loop in CPMV to be removed by digestion with *Nhe*I and *Aat*II. The sequence can then be replaced by any sequence with *Nhe*I- and *Aat*II-compatible ends.

Construction of pMT7-FMDV-II, pMT7-HIV and pMT7-HRV

Three different sequences were designed to be substituted for the sequence between the N*he*I and A*at*II sites of the mutated M RNA sequence. In all three cases the foreign sequences substituted for wild-type sequences encoding six amino acids. The first sequence to be substituted into VP23 consisted of oligonucleotides encoding residues 735-752 from the transmembrane

glycoprotein gp41 from human immunodeficiency virus (HIV-1). This sequence was elected because a synthetic peptide for this region is recognised in enzyme-linked immunosorbent assays (ELISA) by antisera from seropositive AIDS patients and is capable of including antibodies which neutralise a range of HIV-1 isolates. The second sequence consists of the nucleotide sequence encoding residues 85-99 from VP1 of human rhinovirus 14 (HRV-14). In both cases, the oligonucleotides were designed to contain restriction enzyme sites to facilitate screening. The sequences of the oligonucleotides and the effect of the substitutions on the acid sequence of VP23 are shown in Figures 5 and 6. The methods used for the construction of pMT7-HIV and pMT7-HRV are given in WO 92/18618.

The third sequence consisted of nucleotides encoding residues 141-160 from VP1 of FMDV serotype O_1. The effect of the substitution on the amino acid sequence of VP23 is shown in Figure 7. The method used for the construction of pMT7-FMDV-II is given in Usha, *et al.* (1993).

The properties of pMT7-HIV and pMT7-FMDV-II transcripts are described in WO 92/18618 and Usha, *et al.* (1993) respectively. pMT7-HIV transcripts, when mixed with pBT7-123 transcripts can replicate in cowpea protoplasts and the resultant modified coat protein can assemble into chimaeric virus particles. Similarly, pMT7-FMDV-II transcripts can replicate in cowpea protoplasts but progeny RNA accumulated at a considerably lower level than that from pMT7-FMDV-I, or from pMT7-601 which contains the wild-type VP23 sequence. No virus particles could be detected in protoplasts infected with pMT7-FMDV-II. The ability of transcripts derived from pMT7-FMDV-II to multiply in whole cowpea plants was also studied by Usha, *et al.* (1993).

No symptoms developed on inoculated plants and no progeny could be detected in either the inoculated or the upper leaves. The reduced infectivity of pMT7-FMDV-II may be attributed to the resultant chimaeric virus particles lacking an amino acid sequence which is present in the wild type virus, and in chimaeric virus produced from pMT7-FMDV-I infections, and is important for virus replication and spread in the plant.

Example 1

Construction of pMT7-FMDV-V, PMT7-HRV-II and pMT7-HIV-III

The small lesion phenotype of pMT7-FMDV-1 and its competitive disadvantage in comparison with wild-type CPMV, suggest that the heterologous sequence may have been inserted at a sub-optimal site. Detailed examination of the 3-D structure of CPMV revealed that proline 23 (Pro^{23}), which lies in the centre of the ßB-ßC loop of the S protein, is particularly exposed on the virus surface and is potentially the optimum site for any insertion. To make use of this fact and to prevent the introduction of repeated sequences which may facilitate reversion, pairs of complementary oligonucleotides were synthesised in which the sequence encoding the heterologous amino acids are flanked by sequences present in wild type CPMV such that the insert is made immediately preceding Pro^{23}. The oligonucleotides terminate in *Nhe*I and *Aat*II compatible ends enabling them to be inserted between the *Nhe*I and *Aat*II sites of either pMT7-FMDV-II (Usha, *et al.* 1993) or its derivative pMT7-FMDV-II, *Aat*II in place of the original FMDV-specific inserts. Such a strategy not only ensures that the heterologous sequences are inserted at the optimal site and that the inserts are not flanked by direct repeats but also ensures that no CPMV-specific sequences are deleted, a fact believed to be important in enabling virus particles to assemble (Usha, *et al.* 1993). The sequences inserted in this manner consisted of residues 141-159 of VP1 of FMDV serotype O_1 (a slightly shorter version of the epitope in pMT7-FMDV-I), residues 85-98 of VP1 of HRV-14 which make up the immunodominant site, *Nlm*-IA [Sherry, *et al.*, J. Virology (1986) 57, 246-257], and an epitope comprising residues 731 -752 from gp41 of HIV-1, the so-called "Kennedy epitope" [Kennedy, *et al.*, Science (1986) 231, 1556-1559]. The sequence of the oligonucleotides used in the constructions is shown in Figure 8. The resulting constructs were designated pMT7-FMDV-V, pMT7-HRV-II and pMT7-HIV-III, respectively.

The construction and properties of plasmids pBT7-123, pMT7-FMDV-I and pMT7-FMDV-II have been described previously (Usha, *et al.* 1993). These constructs and their derivatives were propagated in *Escherichia coli* strain JM83. Oligonucleotides were synthesised on a Pharmacia Gene Assembler Plus synthesiser. Sequence analysis was performed by "dideoxy" method using either *E. coli* DNA polymerase I (Klenow fragment) or Sequenase™ version 2.0.

To construct pMT7-FMDV-V, pMT7-FMDV-II was digested to completion with *Nhe*I (which cleaves at position 2708 of the M RNA sequence) and partially with *Aat*II which cuts once within M RNA sequence (position 2735) and once in the pUC-derived portion of the plasmid position 2617 on the pUC19 map). A pair of complementary oligonucleotides encoding residues 141-159 from VP1 of FMDV serotype 0_1 flanked by sequences encoding residues 18-22 and 23-26 of the CPMV VP23 protein (Figure 8) were phosphorylated, annealed and ligated into N*he*I/A*at*II-digested pMT7-FMDV-II. Recombinants having the desired structure were identified by restriction enzyme mapping and sequence analysis.

To construct pMT7-HRV-II and pMT7-HIV-III, pMT7-FMDV-II was initially partially digested with *Aat*II and their full-length linear molecules recovered after agarose gel electrophoresis. The linearised plasmid was treated with *E. coli* DNA polymerase I (K1enow fragment) to remove the 3' overhangs left by *Aat*II, recircularised and transformed back into *E. coli* strain JM83. A recombinant, designated pMT7-FMDV-*Aat*II, in which the *Aat*II site in the pUC portion of the plasmid had been destroyed but which retained the *Aat*II site in the M RNA specific portion was identified by restriction enzyme analysis. Complementary oligonucleotides encoding residues 85-98 of VP1 of HRV-14 or residues 731-752 of gp41 of HIV-1, flanked by the appropriate CPMV-specific sequences were phosphorylated, annealed and ligated into *Nhe*II/*Aat*II-digested pMT7-FMDV-*Aat*II giving rise to pMT7-HRV-II and pMT7-HIV-III, respectively.

Example 2

Ability of pMT7-FMDV-V, pMT7-HRV-II and pMT7-HIV-III to replicate in whole cowpea plants

When RNA was transcribed from the chimaeric plasmids, mixed with transcripts from pBT7-123 and inoculated on to cowpea plants, in each case the inoculated leaves developed chlorotic lesions typical of a wild-type CPMV infection. RNA hybridisation analysis revealed the presence of M RNA-specific sequences within these leaves. In all three cases the infection could be mechanically transmitted to further healthy cowpea plants. In the case of pMT7-HRV-II and

pMT7-HIV-III the infection spread to the upper leaves of most of the infected plants giving a typical systemic mosaic. However, the infection induced by pMT7-FMDV remained associated exclusively with the inoculated leaves, no systemic symptoms being observed and no viral-specific RNA being detected in the upper leaves of the plants.

When total RNA was extracted from leaves inoculated with pMT7-FMDV-V transcripts and analysed by RT-PCR, only a single band corresponding in size to the product derived from RNA retaining the insert was observed. Even after up to three serial passages, a similar result was obtained. To confirm that the insert had been retained, the products derived from samples taken from plants after initial inoculation and after three serial passages were cloned into bacteriophage M13 and the nucleotide sequence of a representative number of clones determined. All the clones in both instances contained the sequence corresponding to viral RNA which retained the inserted sequence intact. These results indicate that reversion of pMT7-FMDV-V RNA had not occurred at a detectable frequency. Analysis of RNA extracted from purified pMT7-HRV-II and pMT7-HIV-III particles (see below) supported the conclusion that the new constructs are genetically stable, no evidence of reversion being found after 10 serial passages.

Virus particles could be prepared from leaf tissue with either pMT7-HRV-II or pMT7-HIV-III using the standard CPMV purification protocol [van Kammen and de Jager (1978), Cowpea mosaic virus. *CMI/AAB Descriptions of Plant Viruses,* **197**], the yields obtained (1.2-1.5 mg of virus per gram of fresh tissue) being similar to that obtained with wild-type CPMV. By contrast no particles derived from pMT7-FMDV-V could be obtained using the standard procedure or a number of variants of it. This failure was not due to the absence of particles within the infected tissue since large numbers of such particles could be seen by immunosorbent electron micoscopy (ISEM) of tissue homogenate using grids coated with anti-CPMV serum.

Example 3

Immunological properties of chimaeric virus particles derived from pMT7-HRV-II and pMT7- HIV-III

To confirm that the purified pMT7-HRV-II particles possessed the antigenic properties of the inserted sequence, samples of the purified virions were subjected to western blot analysis using a polyclonal antiserum raised against HRV-14. A product corresponding in size to the modified VP23 protein could be detected, confirming the antigenicity of the inserted sequence. No reaction could be seen with the antiserum when samples of wild-type CPMV were analysed in the same way.

When a sample of denatured virus was examined by electrophoresis on a SDS-polyacrylamide gel, only three bands were seen. The largest polypeptide (L) corresponds to the large (VP37) viral coat protein and comigrated with the L polypeptide from wild-type CPMV. The middle band (S_s) corresponds to the small (VP23) viral coat protein harbouring the HIV-1-specific epitope. The fastest migrating band (S^1) represents the C-terminal 192 amino acids of the VP23 protein. Terminal sequence analysis showed that it was derived from the VP23 protein by proteolytic cleavage between the two C-terminal amino acid residues of the insert. Thus S_s, but not S^1, contains the insert and as predicted reacts with gp41-specific antibody by Western blotting. The predicted N-terminal cleavage product consists of only 43 residues and could not be resolved on the gel system used. Both elements of S remain associated with the virion. Because a certain amount of S^1 protein was always present in preparations of CPMV-HIV regardless of how quickly the virus was purified, it is possible that this cleavage occurs *in planta*.

The strategy designed to overcome the limitations of pMT7-FMDV-I has proved to be successful since all three of the new chimaera (pMT7-FMDV-V, pMT7-HRV-II land pMT7-HIV-III) gave wild-type symptoms on the inoculated leaves and showed no sign of reversion. Furthermore, two chimaera grew as well as wild-type CPMV and could be readily purified. The fact that pMT7-FMDV-V gives wild-type lesions on inoculated leaves but fails to spread systemically suggests that these chimaeric virus particles are fully competent for cell-to-cell movement but deficient for long-distant transport. This phenomenon may be related to the observation that particles from pMT7-FMDV-V appear to aggregate into intracellular crystalline arrays making purification problematic. These features are not a result of the length of heterologous sequence since pMT7-FMDV-V contains an insert intermediate in size (19 residues) between those contained in pMT7-HRV-II (14 residues) and pMT7-HIV-III (22 residues).

Example 4

Use of chimaeric pMT7-HRV-II virus particles to raise antibodies to HRV

Particles of pMT7-HRV-II and wild-type CPMV were purified as described in Example 2, injected into rabbits, the antisera collected and used to probe western blots of denatured HRV-14 virus particles. A single band corresponding to VP1 of HRV-14 could be detected using the antiserum raised against pMT7-HRV-II virus particles even when the serum was diluted 1:16,000. No reaction could be seen with the other HRV-14 coat proteins (VP2 and VP3). No reaction with any HRV-14 protein was found when serum raised against wild-type CPMV was used to probe the blots. The ability of pMT7-HRV-II virions to raise antibodies which recognise VP1 of HRV-14 shows that epitopes presented on the surface of CPMV particles are immunogenic.

Example 5

Use of chimaeric pMT7-HIV-III virus particles to raise neutralizing antibodies to HIV

Transcripts derived from pMT7-HIV-III were mixed with transcripts derived from plasmid pBT7-123 and inoculated onto the leaves of 10 day-old cowpea plants. To obtain large yields of recombinant virus particles, samples of leaf tissue showing symptoms characteristic of a CPMV infection were homogenised in 100mM sodium phosphate pH7.0, centrifuged briefly and the supernatant used to inoculate healthy cowpea plants. The plants were harvested 2-3 weeks post-inoculation and chimaeric virus particles purified as described in Example 2. The purified chimaeric virus, designated CPMV-HIV-I, was stored at 4 °C in 10mM sodium phosphate pH7.0 in the presence of 0.05% (w/v) sodium azide. The quality of the preparation was monitored by electron microscopy and by electrophoresis of portions of denatured virus on 15% polyacrylamide/SDS/reducing gels. The proteins were visualised by staining the gel with coomassie brilliant blue R250. Prior to injection into mice the virus preparation was dialysed against phosphate-buffered saline and protein concentration determined by Bio-Rad™ assay.

Adult C57/BL6 mice were immunized at 8 weeks of age. Virus (CPMV-HIV-I or CPMV) was mixed with aluminium hydroxide adjuvant at a ratio of 1:5 with stirring for 30 min at room temperature. Mice (6 per group) were immunized subcutaneously at the back of the neck in 5 sites with a total of 100 µl of virus-adjuvant mixture containing 100 µg virus. At the required intervals animals were bled from the tail, and serum stored at -20°C. All sera were heated at 56°C for 30 min before being assayed for neutralizing antibody.

Mice were given two injections of CPMV-HIV-1 or wild type CPMV at 0 and 35 days and bled from the tail 14 days later. Individual neutralizing titres of HIV-1 IIIB were determined as follows. Dilutions of heat-treated serum were incubated with about 2000 syncytium-forming units (sfu) per ml of virus for 1 h at 37%. Semiconfluent monolayers of C8166 cells (5 x 10^4 cells/well) were prepared in 96-well tissue culture plates, which had been pretreated with poly-L-lysine. Medium was removed and the cells received 50µl of inoculum. These were incubated for 1 h at 37°C before fresh medium was added. Incubation was continued for 3 days at 37°C, and syncytia were counted with the aid of a microscope and the percentage inhibition calculated for each well.

Figure 9 shows that neutralising antibody was produced in mice immunised with CPMV-HIV-I with a 50% neutralizing titre of about 1/1000 in 83% of the mice. Antiserum diluted 1/100 gave a mean neutralisation titre of 97% with 100% mice responding. The response was highly uniform. Figure 9 also shows that a control group injected in parallel with wild type CPMV also gave a neutralizing response to HIV-I. The neutralisation titre was about 10-fold lower than with CPMV-HIV-I with a 50% neutralization titre of about 1/100. This was evidently a *de novo* antibody response as there was no significant neutralization with serum from non-immunized litter mates even at a dilution of 1/10.

The stability of the neutralizing antibody response to the CPMV-HIV chimaera was investigated by bleeding the mice again at 48 days after the second injection. These antisera had no significant neutralizing titre at a dilution of 1/10, indicating that the level of neutralizing antibody had declined by over 100-fold. The neutralizing activity stimulated by wild type CPMV was now also

undetectable.

The same mice were given a third injection of CPMV-HIV-I as before, 3 months after the second injection, and bled 14 days later. The mean neutralization titre was 54% at a 1/1000 dilution with all mice now responding. There was no neutralization at this dilution with serum from mice boosted with wild type CPMV. Thus there was little overall increase in neutralising activity. The stability of the neutralizing antibody response to CPMV-HIV-I was checked with antiserum obtained after 56 days. The titres had fallen but anitsera from all mice still gave significant neutralization at a dilution of 1/10, with a mean value of 58%. Neutralization by antisera from controls inoculated with wild-type CPMV was barely significant.

Neutralization of the homologous HIV-1 strain by antiserum obtained after the third injection was compared with the neutralization of HIV-1 strains RF and SF2. At a dilution at which strain IIIB was neutralized by 92%, RF was neutralized by 78%, and SF2 by 66%. Antisera made against wild type CPMV also neutralised all three strains, but relative to strain IIIB these antisera neutralised RF and SF2 less than antisera raised against CPMV-HIV-I.

It was confirmed that the neutralising antibodies in the antiserum made against the wild type CPMV were all specific for CPMV epitopes, with none made against the HIV-1 gp41 peptide, by antisera with purified wild type CPMV. Three successive adsorptions with CPMV did not significantly reduce the HIV-I neutralizing titre of the anti-CPMV-HIV-1 serum but reduced the neutralising titre of the anti-CPMV serum 5-fold. Thus we conclude that the majority of neutralising antibodies made against the CPMV-HIV chimaera were made against HIV-specific epitopes, but that CPMV stimulates antibodies that cross-react with neutralising epitopes of HIV-I.

Example 6

Construction of cDNA clones of CPMV RNA M and B which can be used to directly inoculate plants.

cDNA clones of RNAs M and B of CPMV were constructed in pUC18 such that the 5' ends of the viral RNAs are abutted directly to the transcriptional start site of the CaMV 35S promotor. In addition, the RNA B clone (pCP1) can be linearised precisely at the 3' end of the viral sequence by restriction enzyme digestion at a unique *MIu*I site and the RNA M clone (pCP2) can be treated similarly with *Eco*RI. Therefore, after digestion with these enzymes, run-off transcripts can be treated which contain no non-viral sequences at either their 5' or 3' ends.

The clones were constructed as described in Dessens, *et al.* [Journal of General Virology (1993) 73, 889-892]. The CaMV 35S promotor was cloned between the *Hind*III and *Pst*I sites of pUC18, during which process the *Hind*III site was lost. This promotor sequence is flanked by *Sst*II and *Stu*I restriction sites, the latter of which allows blunt-end digestion to expose the transcriptional start site. cDNAs were generated for RNAs M and B and the 5' halves of these were cloned independently by digestion with *Sst*I and *Bam*HI respectively and ligation into a *Stu*I / *Sst*I or *Stu*I / *Bam*HI-cut CaMV 35S vector. The 3' halves of the RNAs were cloned into these constructs by utilising previously constructed cDNA clones (pMT7-601 and pBT7-123) which had been engineered such that the 3' ends could be precisely exposed by restriction enzyme digestion. RNA M was cloned on a *Pst*I / *Eco*RI fragment and RNA B on a *Bgl*II / *Eco*RI fragment.

The CaMV 35S promotor utilises host plant DNA-dependent RNA polymerases and is highly active. Therefore an infection can be generated in the plant host simply by abrading the surface of the primary leaves in the presence of a mixture of linearised pCP1 and pCP2. The host polymerase directs transcription of viral RNA from the CaMV 35S promotor *in vivo*, and in cells where both pCP1 and pCP2 are transcribed, an infection similar to that obtained with wild type CPMV is generated. Therefore an *in vitro* RNA transcription step is no longer required. This represents a considerable advantage over the previous method for inoculating plants both in terms of ease of use and cost.

In order to allow the production of assembled particles of CPMV containing a foreign peptide which has been inserted immediately preceding the Pro^{23} residue in the ßB-ßC loop of the small capsid protein, and in which the corresponding foreign nucleic acid has been inserted into the CPMV genome in the absence of direct sequence repeats flanking the insert and as an addition

to the existing nucleic acid, cDNA clone pCP2 was mutated as described for pMT7 earlier in this specification to create a unique *Aat*II site at position 2735 of the RNA M sequence. This was designated pCP2-*Aat*II.

Oligonucleotide sequences encoding various foreign peptides (see Table 2) were substituted for the sequence between the *Nhe*1 and *Aat*II sites of pCP2-*Aat*II as described in Example 1. The pCP2-*Aat*II variants and pCP-1 were linearised and inoculated onto the primary leaves of cowpea plants. In all cases infections developed and stable chimaeric virus particles expressing the appropriate foreign peptide were recovered from plants.

The following Examples illustrate the application of the present invention to plant viruses other than CPMV. The foreign peptide inserted in the coat protein of the assembled plant virus in question is the MUC1(16) epitope. The MUC1(16) epitope is a 16-mer peptide "GVTSAPDTRPAPGSTA" derived from the extracellular domain, tandem repeat sequence of human polythropine epithelial cell mucin (PEM) [for a review, see Apostolopoulos, V. and McKenzie, I.F.C. (1994) *Crit. Rev. Immunol.*, 14, pp.293-309].

Example 7

Inspection of the crystal structure of Southern bean mosaic virus (SBMV) strain C reveals that a portion of the loop between the βH and βI strands is well exposed upon the surface of the virus at the five-fold and quasi-six fold axes. This portion of the loop comprises amino-acids 251 to 255 of the linear coat protein sequence and nucleotides 3967 to 3981 of the genomic RNA sequence.

The cDNA of the complete 4194 bp RNA genome of SBMV is cloned into a derivative of pBluescript II plasmid vector lacking the T7 and T3 promoters using standard molecular biological techniques. The cDNA is cloned immediately downstream of a bacteriophage T7 such that a unique restriction enzyme site is present at the 3' terminus of the cDNA, thus allowing linearisation of the recombinant plasmid to generate run-off transcripts which mimic the wild-type RNA. As an alternative, the cauliflower mosaic virus (CaMV) 35S promoter may be used.

A sub-clone is then made from this full-length cDNA clone by inserting the *Bgl*II to *Xmn*I fragment (genomic RNA nucleotides 3165 to 4161), which contains within it the whole coat protein open reading frame, into *Bam*HI/*Hinc*II digested pBluescript II. This sub-clone is further manipulated via site-directed mutagenesis at genomic nucleotide positions 3969 (change A to C) and 3984 (G to T) to create *Bam*HI and *Hpa*I restriction sites, respectively.

The modified subclone is digested with these enzymes and separated from the small excised fragment which is replaced by oligonucleotides coding for the excised nucleotide sequence plus nucleotides coding for an epitope sequence MUC1(16). The following five constructs contain the inserted peptide sequence either between coat protein amino-acids 251 and 252, or 252 and 253, or 253 and 254, or 254 and 255, or 255 and 256 (see Fig. 11).

The modified region of the coat protein from each of these constructs is isolated on a *Hind*III/*Avr*II fragment (genomic nucleotides 3434 to 4096) and used to replace the corresponding fragment in the full-length cDNA clone of the virus. Each of these clones is then linearised at the 3' terminus of the cDNA and, in the case of a T7 bacteriophage promoter construct, used to generate capped run-off RNA transcripts which are then inoculated onto the host-plant (*Vigna unguiculata*), or inoculated directly when under the control of the 35S promoter.

The inoculated plants are monitored for symptoms, and the strength of symptoms, yield and stability for each construct are assessed in order to determine the optimal insertion site. If desirable, purified virus may also be used to immunise experimental animals in order to determine the levels of immune response generated by each construct.

This Example can be adapted to allow insertion in any of the exposed loops of SBMV. Similarly, any peptide epitope sequence can be used instead of MUC1(16).

Example 8

This example describes the determination of an insertion site for epitopes by alignment of the primary sequence of a virus whose structure is unknown (lucerne transient streak virus, LTSV), against those of viruses whose structure has been determined.

The crystal structures of two sobemoviruses, SBMV and Sesbania mosaic virus (SMV), have been solved at high resolution. Comparison of the crystal structures reveals that all the secondary structural elements are well conserved between the viruses and, in particular, the protruding loop between the βH and βI is almost identical in shape and location between the two viruses. This structural element would therefore be expected to be well conserved in all sobemoviruses.

Alignment of the primary sequences of LTSV, SBMV and SMV shows a strong conservation of residues between the three viruses within the βH strand region and significant sequence homology within the βI strand (see Fig. 12). This allows the loop region of LTSV to be inferred as spanning amino acids 218 to 224 of the coat protein.

The 4.275 kb RNA genome is cloned as cDNA, as described for SBMV in Example 7. The genomic clone is then modified by site directed mutagenesis at position 3959 (C to T) and position 3998 (T to C) to create unique *Pst*I and *Kpn*I restriction enzyme sites, respectively. The modified genomic clone is digested with these restriction enzymes and separated from the small excised fragment which is replaced by oligonucleotides coding for the excised nucleotide sequence plus nucleotides coding for the epitope sequence MUC1(16). The following six constructs contain the epitope sequence either between coat protein amino acids 218 and 219, or 219 and 220, or 220 and 221, or 221 and 222, or 221 and 223, or 223 and 224 (see Fig. 13).

Each of these clones is linearised at the 3' terminus of the cDNA and, in the case of a T7 bacteriophage promoter construct, used to generate capped run-off RNA transcripts which are then inoculated onto the host-plant (*Nicotiana clevelandii*), or inoculated directly when under the control of the 35S promoter.

The inoculated plants are monitored for symptoms, and the strength of symptoms, yield and stability for each construct are assessed in order to determine the optimal insertion site. If

desirable, purified virus may also be used to immunise experimental animals in order to determine the levels of immune response generated by each construct.

This Example can be adapted to allow insertions in any of the exposed loops of LTSV. Similarly, any peptide epitope sequence can be used instead of MUC1(16).

Example 9

This example describes the determination of epitope insertion sites in a virus (red clover necrotic mosaic virus, RCNMV) whose coat-protein morphology is known to be similar to a second virus (tomato bushy stunt virus, TBSV) belonging to a different virus family whose crystal structure has been solved. In this case, there is only minimal homology between the two coat protein primary sequences, hence secondary structure prediction algorithms are used to assist in the assignment of a particular loop region.

The crystal structure of the coat protein of TBSV reveals that each of the 180 coat protein subunits forming the T=3 icosahedron consists of two β-barrel domains. The first domain forms the surface of the virus particle and is termed the S domain and is equivalent to the single domain found in SBMV. The second, much smaller, domain forms a surface protrusion at right angles to the plane of the S domain. This P domain forms dimeric interactions with the P domain of a neighbouring coat-protein subunits at the strict and quasi two-fold axes of the icosahedron. The presence of the P domain causes the virions to appear distinctly granular when examined under the electron microscope. Between the S and P domains is a short flexible linker followed by a pair of β-strands connected by a loop which appears to be highly exposed on the viral surface with no obvious role in the contacts between subunits. This loop provides a target for epitope insertions.

Dianthoviruses (e.g. RCNMV) also appear distinctly granular when subjected to electron microscopy, and this together with the size of the coat proteins and their limited homology with those of tombusviruses suggests that they may have structural similarity. Alignment of the coat protein sequences of RCNMV and TBSV (see Fig. 14) using a Lipman-Pearson alignment

algorithm, which recognises sequence conservation as well as identity, gives a similarity index of 26.9 (strict homology is 23%). From the alignment it can be seen that the S domain is better conserved (TBSV residues 100 to 269, strict homology 36%) than the P domain (TBSV residues 270-388) which is poorly conserved.

The loop of interest comprises TBSV residues L^{280} A^{281} G^{282} and the sequence around this region shows some conservation in RCNMV, however secondary structure prediction algorithms are also used to predict the location of β-strands and hence the loops which lie between them. Fig. 15 shows a Chou-Fasman β-region prediction plot of RCNMV residues 214-257 using an algorithm based upon the structures found in 64 proteins. This type of plot is claimed to be 80% accurate at predicting beta strands of interest. The plot suggests that beta strands of interest are located between residues 214-221 and 226-228 , hence the loop at the tip of the domain will be residues 222-225 and residues 245-248. A more sophisticated prediction algorithm, the EMBL PHDsec program based upon trained neural networks, may also be used. The resulting output for the region of interest is shown in Fig. 16. This locates the β-strands to residues 220-223 and 227-239, therefore the loop is comprised of residues 224-226. Combining the two sets of data, the loop will lie within the region spanned by residues 222 to 226.

Dianthoviruses have a bipartite RNA genome, both RNAs being required for infectivity. Accordingly, RCNMV RNAs 1 and 2 are cloned as cDNA, using standard molecular biological techniques, into a suitable vector, downstream of a CaMV 35S promoter. As an alternative, the T7 promoter may be used. Both clones are engineered such that they can be linearised at the 3' termini of the cDNAs.

The cDNA genomic clone of RNA1 is modified by site directed mutagenesis at positions 3078 (A to G) and 3081 (G to C), to create a unique *Apa*LI restriction site, and at positions 3108 to 3111 (ACTC to GTTA) to create a unique *Hpa*I restriction site. The mutation at position 3081 is not silent, however the correct codon is restored when ligating in oligonucleotides to generate the epitope insertion. The modified genomic clone is digested with these restriction enzymes and separated from the small excised fragment which is replaced by oligonucleotides coding for the excised nucleotide sequence plus nucleotides coding for the epitope sequence MUC1(16). The

following six constructs, contain the epitope sequence either between coat protein amino acids 221 and 222, or 222 and 223, or 223 and 224, or 224 and 225, or 225 and 226, or 226 and 227 (see Fig. 17).

Each of these clones is linearised at the 3' terminus of the cDNA and, in the case of a T7 bacteriophage promoter construct, used to generate capped run-off RNA transcripts. These are then mixed with similar transcripts from the linearised cDNA clone of genomic RNA2 and inoculated onto the host-plant (*Nicotiana clevelandii*). Linearised clones are inoculated directly when under the control of the CaMV 35S promoter.

The inoculated plants are monitored for symptoms, and the strength of symptoms, yield and stability for each construct are assessed in order to determine the optimal insertion site. Purified virus may also be used to immunise experimental animals in order to determine the levels of immune response generated by each construct.

This Example can be adapted to allow insertions in any exposed loops of RCNMV. Similarly, any peptide epitope sequence can be used instead of MUC1(16).

The following Examples demonstrate application of the present invention to non-icosahedral plant viruses.

Example 10 : Tobacco mosaic virus

Tobacco mosaic virus (TMV) is a member of the tobamovirus genus of plant viruses whose particle structure is that of a rigid rod 300nm long and 18 nm in diameter with a central hole of 4nm. The rod is formed by 2100 coat protein subunits arranged in a right-handed helix with a pitch of 2.3 nm and 16.33 subunits per turn which encapsidate the viral RNA of 6.5 Kb The structure of the individual coat protein subunits of this virus has been solved at atomic resolution and consists of a core made up of a right-handed, four anti-parallel α-helix bundle with left and right slewed helices (LS & RS) and left and right radial helices (LR & RR) extending outwards from an RNA binding site (Namba, *et al.* 1989). The crystal structure suggests that possible insertion sites for peptides or polypeptides could be at or close to the termini (both N- and C-)

and in the loop connecting the RS and RR helices. In fact, expression systems have been developed which exploit two of these potential sites.

Early work on direct fusions to the C-terminus demonstrated that whilst some short additional sequences were allowed (7 amino-acids) others prevented virus assembly and movement through the plant (Takamatsu, *et al.* 1990). However it was subsequently found that C-terminal extensions of TMV could be exploited as an efficient expression system provided that mechanisms to allow for the production of unmodified coat protein as well as extended coat protein were employed. It has been demonstrated that direct fusions to the C-terminus can be achieved by utilizing the phenomenon of stop codon read-through (Hamamoto, *et al.* 1993; Sugiyama, *et al.* 1995; Turpin, *et al.* 1995) whereby approximately 5% of the coat protein subunits possess C-terminal extensions. A second method for generating a mixture of wild-type coat protein and modified coat protein is to inoculate plants with a mixture of two viruses, one being a complete virus genome containing a modified coat protein and the second being a virus lacking the movement protein but containing an unmodified coat protein (Beachy, *et al.* 1996). It has also been shown that viruses with additions to the C-termini of all the coat protein subunits can assemble and move systemically in infected plants and are stable over multiple passages, however in this case the additional sequences are not fused directly to the C-terminus of the coat protein, but are inserted at a site 5 or 6 amino acids from the C-terminus (Fitchen, *et al.* 1995; Beachy, *et al.* 1996).

Short peptides (4 or 12 amino-acids) have also been inserted in the loop connecting the RS and RR helices, between coat protein amino-acids 63 and 66 accompanied by a deletion of amino-acids 64 and 65 (Turpen, *et al.* 1995) . These modified viruses assemble and systemically infect inoculated plants and can be recovered with high yields.

Example 11: Potato Virus X

Potato virus X (PVX) is a member of the potexvirus genus of plant viruses whose particle structure is that of a flexuous rod 515nm long and 13 nm in diameter with a central hole of 1.7nm. The rod is formed by 1270 coat protein subunits arranged in a right handed helix with a

pitch of 3.4 nm and 8.875 subunits per turn which encapsidate the viral RNA of 6.4kb. A model of the structure of PVX has been proposed based upon tritium planigraphy and immunological analysis (Baratova, *et al.* 1992a , 1992b) where the N-terminus of the coat protein is exposed between amino-acid residues 1-33 and forms a beta sheet composed of three beta strands. In contrast to TMV, the C-terminus is not exposed and it is thought to lie underneath the surface structure formed by the N-terminus. Therefore, based upon this low resolution data, fusions of peptides and polypeptides to the N-terminus should be possible.

It has been demonstrated that polypeptides as large as 27 kDa can be fused to the N-terminus of the PVX coat protein (Santa-Cruz, *et al.* 1996), however, as was demonstrated for C-terminal fusions to TMV, it proved necessary to engineer the construct in such a way that a mixture of free coat protein and polypeptide/coat protein fusion was produced in order to allow virus assembly and spread within the plant. This was achieved by inserting a 16 amino-acid sequence between the final codon of the polypeptide and Pro^4 of the coat protein. This sequence codes for the core FMDV 2A protease which autocatalytically cleaves between its final amino-acid and the proceeding proline residue of the PVX coat protein, however this autocatalytic event occurs inefficiently such that the majority of protein produced is the fusion product with a smaller proportion of free polypeptide and coat protein, however the amount of free coat protein generated is sufficient to ensure efficient assembly and spread of virus particles which also carry the polypeptide-coat protein fusion.

Example 12

Preparation of PVX-MAST8

In order to utilise linearised DNA for the inoculation of plants, the full length viral cDNA from pTXS.L2a-CP (patent application No. 9420989.7) was transferred into plasmid pCP8 such that the viral cDNA was fused directly to the transcriptional start site of the CaMV 35S promoter. This plasmid was called pTXS6. Following the insertion of peptides, as described below, the DNA was linearised with SpeI and used to inoculate young *Nicotiana benthamiana* leaves in order to produce viral particles (Brennan et al, 1999).

Plasmid pTXS6 was digested with *Nhe*I and *Afl*II and five oligonucleotides coding for the excised

regions of the PVX vector, an initiating methionine and amino acids 1-38 (GQNNGNQSFEEDTEKDKPKYEQGGNIIDIDF) from the D2 peptide derived from an *S. aureus* fibronectin binding protein (FnBP) were inserted (as illustrated below). This plasmid, termed pTXS6-MAST8, therefore contained the coding sequence for amino acids 1-38 in frame and directly in front of the 16aa sequence coding for the core FMDV 2A protease.

NheI, ATG, D2 peptide, AflII

CTAGCATG.....AATTTTGACCTTC^

GTAC.....TTAAAACTGGAAGAATT^

The DNA was linearised and used to inoculate *Nicotiana benthamiana* using 0.5 µg of linear DNA per leaf. After 14 days, the leaves were harvested and virus particles extracted by homogenisation in 50 mM sodium borate pH 9, 10 mM β-mercaptoethanol. Following three chloroform extractions, the virus particles were precipitated once with 5% (w/v) PEG 8000 and three times with 5% PEG 8000, 0.25M NaCl. The final pellet was resuspended in 50 mM sodium borate pH 9, 20% glycerol and stored at –70°C. The final concentation and quality of these viral particles, termed PVX-MAST8, was determined by comparison with PVX standards purified using standard protocols on SDS PAGE gels. From this it was determined that approximately 10% of the viral coat protein subunits carried the peptide fusion.

Mice were immunised subcutaneously with PVX-MAST8 in adjuvent and high titres of FnBP-specific antibody were obtained. This sera was shown to completely inhibit the binding of fibronectin to immobilised recombinant FnBP.

This study demonstrates that the D2 peptide is highly immunogenic when expressed on the surface of PVX and that the viral particles can be purified from plant sap.

Example 13

Preparation of pTXS2-ATZ27

A number of proteins have been expressed on the surface of PVX using the original pTXS.L2a-CP

vector described above. A single-chain antibody fragment against the triazine herbicide atrazine was amplified by PCR, digested with *Cla*I and *Afl*II and ligated into pTXS.L2a-CP digested with the same enzymes. Following sequence analysis to confirm that the scFv was in frame with the 2A protease of FMDV and the viral coat protein, the plasmid DNA (pTXS2-ATZ27) was linearised with *Sph*I and infectious transcripts produced using standard protocols. The infectious transcripts were then used to inoculate *Nicotiana benthamiana* plants.

Fourteen days after inoculation, total soluble protein was extracted from a systemically infected leaf and analysed by western blot analysis, using antisera raised against the coat protein of PVX. Both the viral fusion protein of 53kDa and the cleaved viral coat protein of 24.8 kDa were clearly detected. The number of viral coat protein subunits carrying the peptide fusion ranged from 5-50% in different plants.

In order to confirm affinity for atrazine, ELISAs were carried out using atrazine-BSA coated plates. Plant sap was extracted from systemically infected leaves, using 1 x PBS buffer and the amount of total soluble protein estimated using a BIO-RAD protein assay. Total soluble protein at 50-0.1 mg/ml was then added to the ELISA plate. In a similar manner, the amount of PVX in each sample was estimated and plant extracts containing 50-0.1 mg/ml of PVX were added to the ELISA plate. In each case the viral fusion protein was then detected with an anti-PVX polyclonal antisera. Affinity for the atrazine was clearly demonstrated in those samples expressing the viral fusion protein compared with wild type PVX infected or non-infected plant tissue. In each case an increase in relative affinity correlated with an increase in the amount of viral coat protein subunits carrying the peptide fusion.

Example 14: Plum Pox Virus

Plum Pox virus (PPV) is a member of the potyvirus genus of plant viruses (family *potyviridae*) whose particle structure is that of a flexuous rod 750nm long and 11-13 nm in diameter. The rod is formed by more than 2000 coat protein subunits arranged in a right handed helix with a pitch of 3.4 nm which encapsidate the viral RNA of 9.8kb. Very little structural information is known about potyviruses, however immunological and protease treatment studies have shown that both

the N- and C-termini are surface exposed (Shukla, *et al.* 1988; Shukla and Ward, 1989) and might therefore be useful sites for the insertion of peptides or polypeptides. In addition there are naturally occurring mutants of PPV which are non-aphid transmissible which have a deletion of 15 amino-acids within the N-terminus of the coat protein (NAT mutants: Maiss, *et al.* 1989) which includes the Gly of the Asp-Ala-Gly amino-acid triad essential for aphid transmission (Atreya ,*et al.* 1990 and 1991). Therefore PPV is naturally capable of possessing at least an extra 15 amino-acids within its N-terminus and the site of the naturally occurring deletion is a logical place to insert peptides or polypeptides. In addition, the length of the N-terminus is highly variable between potyviruses, suggesting that potyviruses will tolerate different sequence lengths, and this region of the coat protein is known to be highly immunogenic.

Insertions of 15 and 30 amino-acids have been made within the N-terminus as described above (Fernandez-Fernandez *et al.* 1998). The chimaeric viruses were infectious, genetically stable and accumulated in the infected plants to wt levels and gave similar yields on purification. In contrast to some of the TMV systems and the PVX system, additions to all the coat protein subunits were tolerated.

Example 15 : Tobacco Rattle Virus

Tobacco rattle virus (TRV) is a member of the tobravirus genus of plant viruses whose particle structure is that of a rigid rod. TRV has two genome components which are encapsidated separately into rods of about 190nm in length and 50 to 115nm in length depending upon the isolate. The rods are 23 nm in diameter with a central hole of 5nm. The coat protein subunits are arranged in a right handed helix with a pitch of 2.5 nm and the number of subunits per turn has been estimated at either 25.33 or 32.33. The longer particle encapsidates genomic RNA1 which is 6.8 kb in length and the shorter particle encapsidates RNA2 which ranges in size from 1.8 kb to 4.5 kb depending upon the isolate. RNA1 is capable of independent replication and systemic spread in infected plants, however no virus particles are formed in these infections since the coat protein gene is carried on RNA2 and transmission is difficult. Infections containing both RNA1 and RNA2 produce virus particles and are readily transmissible between plants by mechanical inoculation and nematodes.

No detailed structural information is known for TRV, however coat protein sequence alignments of tobraviruses with those of tobamoviruses suggest that they share a common evolutionary origin (Goulden, *et al.* 1992) and may therefore have very similar structures. Immunological and proteolytic analysis of the tobraviral particle shows that the C-terminus is exposed upon the viral surface and can be removed without affecting particle integrity (Legorburu, *et al.* 1996) and it has also been shown by NMR spectroscopy that the C-terminus is highly mobile (Mayo, *et al.* 1993). Therefore, as is the case with TMV, the C-terminus of TRV may be a suitable site for the insertion of peptides and polypeptides.

Although TRV and TMV are thought to be structurally very similar, there may be advantages in using the former in preference to the latter for the expression of peptides and polypeptides as follows :

1. TRV does not require coat protein to systemically invade host plants, therefore additions to the C-terminus will not interfere with virus spread.

2. The C-terminus is much longer than that of TMV and is not required for particle stability, and may therefore be replaced by foreign sequences. Replacement is less likely to interfere with virus assembly than addition.

3. The C-terminus is highly mobile and this flexibility may allow the addition of foreign sequences without interfering with viral assembly.

The cDNA of the complete 6791 b RNA1 of TRV strain PSG is synthesized and cloned into a derivative of a pBluescript II SK+ vector (lacking the T7 and T3 promoters) using standard molecular biological techniques. The cDNA is cloned such that the 5' end is at the transcriptional start of a bacteriophage T7 promoter and the 3' end can be exposed by linearisation with a unique restriction enzyme site in order to generate run-off transcripts which mimic the wild-type RNA sequence. As an alternative, the cauliflower mosaic virus (CaMV) 35S promoter is used. The cDNA of the complete 1905 b RNA2 of TRV strain PSG is synthesized and cloned in a similar

fashion. Infections are generated in host plants (*Nicotiana clevelandii*) by linearising the RNA1 and RNA2 clones at their 3' termini and either inoculating with a mixture of run-off transcripts in the case of T7 controlled clones or mixing and inoculating the plasmid DNA directly in the case of 35S controlled clones.

The requirement for C-terminal sequences of the coat protein for assembly is then investigated by stepwise deletions introduced into the coat protein gene on the RNA2 cDNA clone, utilising naturally occurring unique *Sal* I and *Ppu*M I restriction enzyme sites at RNA2 positions 1125 and 1224 respectively (Note : *Sal* I is present in the polylinker of p Bluescript II, but is removed during the cloning of the cDNA of RNA2). The RNA2 clone is digested with *Sal* I and *Ppu*M I and separated from the 100bp excised fragment which is replaced by oligonucleotides containing increasingly long deletions of the coat protein C-terminus. Five constructs are made containing deletions of 5 , 10 , 15 , 20 and 22 amino-acids (Fig. 18). These constructs and the unmodified clone are inoculated onto plants (as described above) and monitored for symptom development. Sap extracts from infected leaves are then examined under the electron microscope for the presence of viral particles and, where particles are present, purifications are done to assess yield.

The largest coat protein deletion which allows particle assembly comparable to undeleted TRV is then selected as a potential carrier for peptides and polypeptides. Hence, overlapping oligonucleotides are assembled containing the sequence corresponding to this deletion plus the coding sequence of choice for a peptide or polypeptide to be inserted between the ultimate amino-acid of the deletion and the stop codon. This construct is inoculated onto plants as described above and the strength of symptoms, virion assembly, yield and stability of the construct is assessed. In the event that particle assembly is prevented by the addition of the non-native sequences then additional constructs are made containing smaller deletions of the C-terminus of the coat protein.

Purified virus from successful constructs may further be used to immunise experimental animals to determine the strength of the immune response generated by the inserted, non-native

sequences.

Table 2

Foreign peptide sequences expressed as chimaeric virus particles produced by direct inoculation of plants with cDNA clones.

CONSTRUCT	LENGTH (amino acids)	PEPTIDE SOURCE
HIV-1	22	amino acids 732-752 of gp41 of HIV-I strain IIB
HIV-3	6	amino acids 312-317 of gp120 of HIV-I strain IIIB (the V3 loop)
HIV-4	11	amino acids 140-150 of gp120 of HIV-I strain IIIB (the V1 loop)
HIV-5	11	amino acids 117-127 of HIV-I strain IIIB
FMDV-5	19	amino acids 141-159 of VP1 of FMDV strain 0_1 (the G-H loop)
FMDV-12	21	A peptide sequence form VP1 of FMDV strain CS8 (the G-H loop)
FMDV-13	23	A peptide sequence from VP-1 of FMDV strain A10 (the G-H loop)
FMDV-14	10	amino acids 40-49 of VP1 of FMSV strain 0_1 (the B-C loop)
PARVO-1	17	amino acids 13-29 of VP2 of canine parvovirus
PARVO-2	17	the insert sequence of PARVO-1 in reverse
PARVO-3	17	as flanking sequence variant of PARVO-1
GNRH-1	10	an immunodominant epitope form pig gonadotrophin releasing hormone
MAST-1	30	derived from the fribronectin binding protein of *Staphylococcus aureus*
MAST-2	38	a longer version of MAST-1
HRV-2	14	amino acids 85-98 of VP1 of HRV strain 14

References :

Atreya C.D., Raccah B. & Pirone T.P., 1990. A point mutation in the coat protein abolishes aphid transmissibilty of a potyvirus. *Virology* **178** : 161-165 .

Atreya P.L., Atreya C.D. & Pirone T.P., 1991. Aminoacid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. *Proc. Natl. Acad. Sci. USA* **88** : 7887-7891 .

Baratova L.A., Grebenshchikov N.I., Shishkov A.V., Kashirin I.A., Radavsky Y.L. Jarvekulg L., & Saarma M., 1992(b). The topography of the surface of potato virus X : tritium planigraphy and immunological analysis. *J. Gen. Virol.* **73** : 229-235.

Baratova L.A., Grebenshchikov N.I., Dobrov E.N., Gedrovich A.V., Kashirin I.A., Shishkov A.V., Efimov A.V., Jarvekulg L., Radavsky Y.L. & Saarma M., 1992(b). The organisation of potato virus X coat proteins in virus particles studied by tritium planigraphy and model building. *Virology* **188** : 175-180.

Beachy R.N., Fitchen J.H. & Hein M.B., 1996. Use of plant viruses for delivery of vaccine epitopes *In : Engineering Plants for Commercial Products and Application. Annals of the New York Academy of Sciences Vol. 792.* Eds. : Collins G.B. & Shelford. New York Academy of Sciences USA.

Fernandez-Fernandez M.R., Martinez-Torrecuadrada J.L., Casal J.I. & Garcia J.A., 1998. Development of an antigen presentation system based on plum pox potyvirus. *FEBS Letters* **427** : 229-235.

Fitchen J., Beachy R.N. & Hein M.B., 1995. Plant virus expressing hybrid coat protein with added murine epitope elicits autoantibody response. *Vaccine* **13** : 1051-1057.

Goulden M.G., Davies J.W., Wood K.R. & Lomonossoff G.P., 1992. Structure of tobraviral particles : a model suggested from sequence conservation in tobraviral and tobamoviral coat

proteins. *J. Mol. Biol.* **227** : 1-8.

Hamamoto H., Sugiyama Y., Nakagawa N., Hashida E., Matsunaga Y., Takemoto S., Watanabe Y. & Okada Y., 1993. A new tobacco mosaic virus vector and its use for the systemic production of angiotensin-I-converting enzyme inhibitor in transgenic tobacco and tomato. *Bio/technology* **11** : 930-932.

Legorburu F.J., Robinson D.J. & Torrance L., 1996. Features on the surface of tobacco rattle tobravirus particle that are antigenic and sensitive to proteolytic digestion. *J. Gen. Virol.* **77** : 855-859.

Maiss E., Timpe U., Brisske A., Jelkmann W., Casper G., Himmler G., Mattanovich D. & Katinger H.W.D., 1989. The complete nucleotide sequence of plum pox virus RNA. *J. Gen. Virol.* **70** : 513-524.

Mayo M.A., Brierly K.M. & Goodman B.A., 1993. Developments in the understanding of the particle structure of tobraviruses. *Biochimie* **75** : 639-644.

Namba K., Pattanayek R. & Stubbs G., 1989. Visualisation of protein-nucleic acid interactions in a virus. Refined structure of intact tobacco mosaic virus at 2.9 angstrom resolution by X-ray fibre diffraction. *J. Mol. Biol.* **208** : 307-325.

Santa-Cruz S., Chapman S., Roberts A.G., Roberts I.M., Prior D.A.M. & Oparka K.J., 1996. Assembly and movement of a plant virus carrying a green fluorescent protein overcoat. *Proc. Natl. Acad. Sci. USA* **93** : 6286-6290.

Shukla D.D., Strike P.M., Tracy S.L., Gough K.H. & Ward C.W., 1988. The N and C termini of the coat proteins of Potyviruses are surface located and the N terminus contains the major virus specific epitopes . *J. Gen. Virol.* **69** : 1497-1508.

Shukla D.D. & Ward C.W., 1989. Structure of potyvirus coat proteins and its application in the taxonomy of the potyvirus group. *Adv. Virus Res.* **36** : 273-314.

Sugiyama Y., Hamamoto H., Takemoto S., Watanabe Y. & Okada Y., 1995. Systemic production of foreign peptides on the particle surface of tobacco mosaic virus. *FEBS Letters* **359** : 247-250.

Takamatsu N., Watanabe Y., Yanagi H., Meshi T., Shiba T. & Okada Y., 1990. Production of enkephalin in tobacco protoplasts using tobacco mosaic virus RNA vector. *FEBS Letters* **269** : 73-76.

Turpen T.H., Reini S.J., Charoenvit Y., Hoffman S.L., Fallarme V. & Grill L.K., 1995. Malarial epitopes expressed on the surface of recombinant tobacco mosaic virus. *Bio/technology* **13** : 53-57.